

WHAT IS CLAIMED IS:

- 1 1. A system for classifying files of non-textual subject data comprising:
2 a system decision module that includes:
3 (a) a task component having a plurality of
4 classification tasks arranged in a sequential progression of decision
5 making, said sequential progression of decision making including a
6 plurality of classification nodes for assigning classes, at least some of
7 said classification nodes including algorithms for determining which of
8 a plurality of alternative next classification nodes is to be encountered
9 in said sequential progression of decision making;
10 (b) an algorithmic component for selecting an
11 algorithm for each of said classification tasks, said algorithm being
12 configured to execute at least one of content-based analysis for
13 processing content-based data and meta-data analysis for processing
14 meta-data;
15 (c) a sub-algorithmic component for selecting at least
16 one sub-algorithmic routine for said algorithm, said sub-algorithmic
17 routine being selected based on said selecting said algorithm; and
18 (d) a learning component for modifying said
19 arrangement of classification tasks according to determinations of
20 the frequencies of assignments of said classes to said files of
21 non-textual subject data.
- 1 2. The system of claim 1 further comprising a system web-service module
2 for providing Internet access to said system decision module.
- 1 3. The system of claim 1 further comprising a system interface module
2 for providing communications among a plurality of system and non-system
3 modules, wherein one of said system modules is said system decision
4 module.
- 1 4. The system of claim 3 wherein each of said non-system modules
2 includes at least one said sub-algorithmic routine.

1 5. The system of claim 3 wherein said system interface module further
2 includes data components for storing data associated with classifying a
3 plurality of said files of said non-textual subject data and at least one control
4 component for executing said sub-algorithmic routines.

1 6. The system of claim 1 further comprising a media input/output module
2 for administering data associated with classifying said non-textual subject
3 data by reading and writing said data among a plurality of modules.

1 7. The system of claim 1 wherein said learning component is configured
2 to identify an algorithm for each of said classification tasks and at least one
3 sub-algorithmic routine for said algorithm.

1 8. The system of claim 1 further comprising a data capturing device
2 configured to capture said content-based data and record said meta-data,
3 said content-based data corresponding to content information of a file of said
4 subject data and said meta-data corresponding to situational environmental
5 data of said data capturing device during a capture of said subject data.

1 9. A method for categorizing files of non-textual data comprising the
2 steps of:
3 establishing a sequential progression of decision making,
4 including using automated processing techniques to define a dependent
5 arrangement of a plurality of task nodes, each said task node being
6 associated with a class for classifying a data file, at least some of said task
7 nodes including algorithms for determining which alternative next task node is
8 to be selected in said sequential progression of decision making, said task
9 nodes including multi-algorithmic task nodes having a plurality of alternative
10 said algorithms for implementing said determination;
11 receiving a file of non-textual subject data; and
12 progressing said file through said sequential progression of
13 decision making, including selecting from among said alternative algorithms
14 at said multi-algorithmic decision nodes at least partially based on prior
15 determinations at previously encountered task nodes in said sequential
16 progression.

1 10. The method of claim 9 wherein said step of establishing includes a
2 learning procedure in which content-based data is extracted from each of a
3 plurality of training images and meta-data is identified for each said training
4 image.

1 11. The method of claim 10 further comprising a step of generating a
2 plurality of learning classes that are descriptive of said training images,
3 including using an association pattern technique, said step of generating
4 including applying content-based analysis for said content-based data and
5 meta-data analysis for said meta-data.

1 12. The method of claim 9 further comprising a step of dynamically modify-
2 ing said sequential progression of decision making, including monitoring said
3 determinations at each of said decision nodes and adjusting for detected
4 patterns in said determinations.

1 13. The method of claim 9 further comprising a step of assigning a
2 semantic description to said file of non-textual subject data for one of
3 organizing said file and matching a query during a search for said file.

1 14. A method for identifying a class for a data file at a classification node
2 comprising the steps of:
3 subjecting an image data file to a transformation function to
4 generate transformed image data, said step of subjecting including transform-
5 ing at least one of content-based data and meta-data, said content-based
6 data corresponding to image data of said file and said meta-data correspond-
7 ing to situationally surrounding conditions of a recording device during a
8 capture of said image data file;
9 performing feature analysis on said transformed image data to
10 derive feature data characteristic of said file; and
11 applying an algorithmic routine utilizing said feature data to
12 generate a class identifiable with said file.

1 15. The method of claim 14 wherein said step of applying includes
2 selecting said algorithmic routine from a plurality of algorithmic routines.

1 16. The method of claim 14 further comprising a step of defining said
2 algorithmic routine for generating said class based on a training procedure by
3 subjecting a plurality of training image data files having characteristics
4 attributable with said class.

1 17. The method of claim 14 wherein said step of applying includes a
2 selection of said algorithmic routine at least partially based on a determination
3 of a previous classification task.

1 18. The method of claim 14 wherein said step of performing said feature
2 analysis includes applying statistical analysis on said transformed image data.